# Assignment 4
# Improving Medical Image Segmentation Models
# UKAN

SATHVIK MERIPE
220010031

March 29, 2025

## Contents

# 1 Introduction

## 1.1 Overview

This assignment focuses on improving the performance of baseline models used for medical image segmentation. The goal is to introduce innovative modifications that go beyond standard data augmentation techniques to enhance segmentation accuracy and robustness.

## 1.2 Objective

The primary objective is to analyze and implement methodologically innovative enhancements to baseline models. The performance of these enhanced models will be evaluated using specific segmentation metrics across three distinct medical image datasets. The aim is to demonstrate measurable improvements in segmentation performance compared to the baseline models.

# 2 Datasets

## 2.1 ISIC 2017 (Skin Lesion Segmentation)

The ISIC 2017 dataset focuses on skin lesion segmentation, providing a collection of dermoscopic images for the detection and segmentation of skin lesions. The dataset includes **2,000** valid image-mask pairs for training and **150** valid image-mask pairs for validation. The directory structure is organized as follows:

- Training images: `ISIC-2017_Training_Data`
- Training masks: `ISIC-2017_Training_Part1_GroundTruth`
- Validation images: `ISIC-2017_Validation_Data`
- Validation masks: `ISIC-2017_Validation_Part1_GroundTruth`

## 2.2 Fetus Head Segmentation

This dataset contains ultrasound images aimed at segmenting the fetus head, which is crucial for monitoring fetal development and health. It includes **799** valid image-mask pairs for training and **200** valid image-mask pairs for validation. The directory structure is organized as follows:

- Training images and masks: `training_set`
- Validation images and masks: `validation_set`

## 2.3 Lumbar Spine Segmentation

The Lumbar Spine Segmentation dataset includes images for segmenting the lumbar spine region, essential for diagnosing spinal conditions. It contains **698** valid image-mask pairs for training and **199** valid image-mask pairs for validation. The directory structure is organized as follows:

- Images: `png/images`
- Masks: `png/masks`

The dataset is further divided into case folders, with images and masks stored separately within each case folder.

# 3 Methodology

## 3.1 Baseline Model (UKAN)

The UKAN model is a U-Net-inspired architecture that integrates Kolmogorov-Arnold Networks (KAN) for enhanced nonlinear approximation capabilities. This model is designed to improve the segmentation of regions of interest in medical images by leveraging the strengths of both CNNs and KANs.

### 3.1.1 Key Components

1. **Encoder-Decoder Framework**
   - Utilizes convolutional layers for downsampling and upsampling.
   - Implements skip connections between the encoder and decoder stages to preserve spatial information.
2. **KAN Integration**
   - Replaces traditional Multi-Layer Perceptron (MLP) blocks with KAN layers, defined as:

     ```
     KANLinear(in_features, hidden_features, grid_size=5, spline_order=3)
     ```

   - Combines spline-based activation functions with depthwise convolutions to enhance feature representation.
3. **Key Innovations**
   - Adaptive activation functions via learnable splines, allowing the model to adapt to the data's complexity.
   - Hybrid architecture that combines the strengths of CNNs with the approximation capabilities of KANs.
   - Position-aware processing through Depthwise Convolution (DWConv) modules, enabling the model to capture spatial hierarchies effectively.

### 3.1.2 Architectural Advantages

- **Better Function Approximation**: KAN layers theoretically outperform MLPs in learning complex patterns, making them suitable for medical imaging tasks that require precise feature localization.
- **Parameter Efficiency**: The use of shared weights in depthwise convolutions reduces the overall model size, making it more efficient.
- **Multi-scale Processing**: Patch embedding layers, implemented using 3×3 convolutions with a stride of 2, enable hierarchical feature learning, capturing both local and global contexts.

The UKAN model shows significant potential for medical imaging tasks, where precise segmentation is crucial. However, its actual performance depends on the specific implementation details and training protocols employed.

## 3.2 Proposed Model (UKAN Upgraded)

The upgraded UKAN model introduces a hybrid CNN-KAN-Transformer architecture designed to enhance segmentation performance in medical imaging tasks. This architecture combines CNNs for local feature extraction, KAN blocks for enhanced nonlinear approximation, and Transformers for global context modeling.

### 3.2.1 Core Innovations

1. **Triple Hybrid Design**
   - The architecture flow is structured as follows:

     ```
     CNN Encoder → KAN Blocks → Transformer Bottleneck → KAN Blocks → CNN Decoder
     ```

2. **CLAHE Integration**
   - Contrast-Limited Adaptive Histogram Equalization (CLAHE) is applied for contrast enhancement, particularly beneficial for medical images (pre processing step).

### 3.2.2 Architectural Breakdown

| **Component** | **Key Features** | **Purpose** |
|---|---|---|
| **CNN Encoder** | 3-stage conv blocks with max pooling | Local feature extraction |
| **KAN Blocks** | Spline-activated linear layers + DWConv | High-accuracy nonlinear mapping |
| **Transformer** | Multi-head self-attention + MLP | Global context modeling |
| **CNN Decoder** | Transposed convs + skip connections | Precise spatial reconstruction |

This architecture represents a significant evolution from standard U-Nets, particularly suited for complex segmentation tasks where both local details and global context are critical. The transformer bottleneck helps overcome traditional CNN's limited receptive field, while KAN blocks provide superior function approximation capabilities. The integration of CLAHE further enhances the model's ability to discern subtle features in medical images by improving contrast.

## 3.3 Training Setup

The training of the UKAN model was conducted using the following configuration:

- **Batch Size**: 8
- **Architecture**: UKAN
- **Input Dimensions**: $256 \times 256$ pixels
- **Loss Function**: BCEDiceLoss
- **Optimizer**: Adam
- **Learning Rate (LR)**: 0.0001
- **Momentum**: 0.9
- **Weight Decay**: 0.0001
- **KAN Learning Rate**: 0.01
- **KAN Weight Decay**: 0.0001
- **Scheduler**: CosineAnnealingLR
- **Number of Workers**: 4

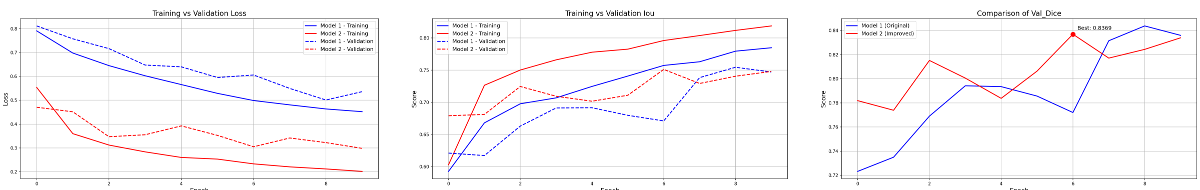
No of epochs ran for each dataset is written in the results section along with metrics.

# 4 Results

## 4.1 ISIC 2017 Dataset

### 4.1.1 Training Progress

The training progress for the models is depicted through the following plots: loss, IoU, and DICE score over 10 epochs.

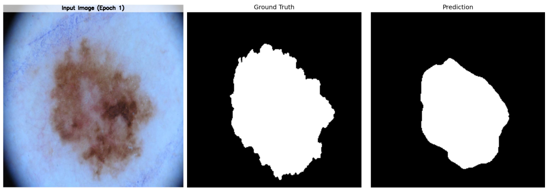



(a) Loss over epochs (b) IoU over epochs (c) DICE score over epochs

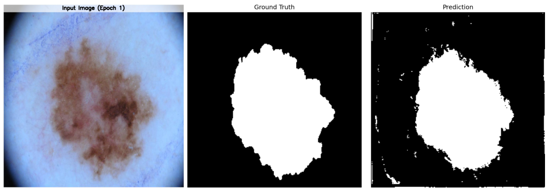
Figure 1: Training progress plots for the ISIC 2017 dataset.

### 4.1.2 Epoch-wise Predictions

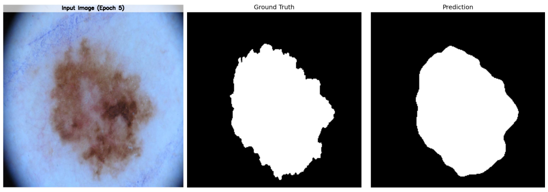
The following images show the model's predictions at different epochs, illustrating the improvement in segmentation accuracy over time. The following images show the model's predictions at different epochs, comparing the normal model (left) with the upgraded model (right).



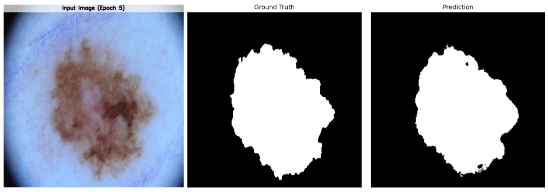

(a) UKAN - Epoch 1



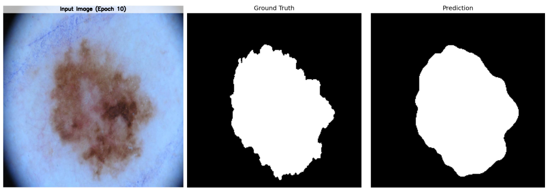

(b) Upgraded UKAN - Epoch 1



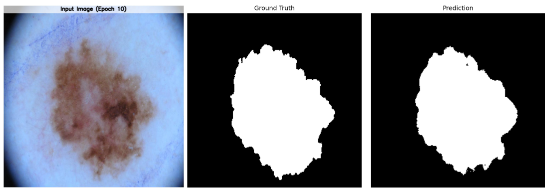

(c) UKAN - Epoch 5




(d) Upgraded UKAN - Epoch 5




(e) UKAN - Epoch 10




(f) Upgraded UKAN - Epoch 10

Figure 2: Comparison of predictions between the normal and upgraded models at different epochs.

### 4.1.3 Model Comparison

Table 1: Model Performance Comparison - ISIC 2017 Dataset

| Metric | Model 1 (Best) | Model 2 (Best) | Improvement | Relative Improvement (%) |
|---|---|---|---|---|
| Loss | 0.4519 | 0.2016 | 0.2504 | 55.40 |
| IoU | 0.7847 | 0.8186 | 0.0339 | 4.32 |
| F1 Score | 0.8794 | 0.9002 | 0.0209 | 2.37 |
| Val Loss | 0.5007 | 0.2982 | 0.2026 | 40.46 |
| Val IoU | 0.7544 | 0.7511 | −0.0033 | −0.44 |
| Val Dice | 0.8438 | 0.8369 | −0.0069 | −0.81 |

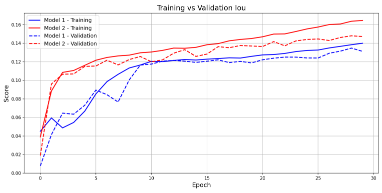
Model 1: UKAN Baseline, Model 2: Upgraded UKAN
Negative values indicate performance degradation

## 4.2 Fetus Head Segmentation Dataset

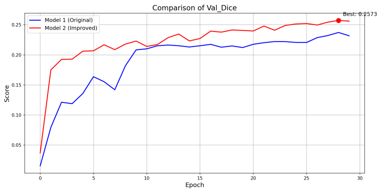
### 4.2.1 Training Progress

The training progress for the models is depicted through the following plots: loss, IoU, and DICE score over 30 epochs.


(a) Loss over epochs




(b) IoU over epochs

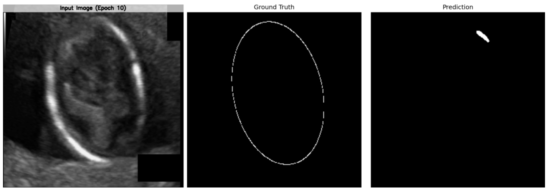



(c) DICE score over epochs

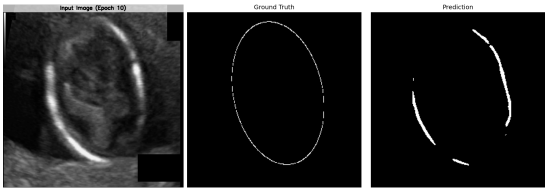
Figure 3: Training progress plots for the Fetus Head Segmentation dataset.

### 4.2.2 Epoch-wise Predictions

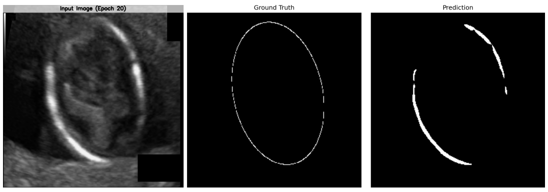
The following images show the model's predictions at different epochs, illustrating the improvement in segmentation accuracy over time. The following images show the model's predictions at different epochs, comparing the normal model (left) with the upgraded model (right).



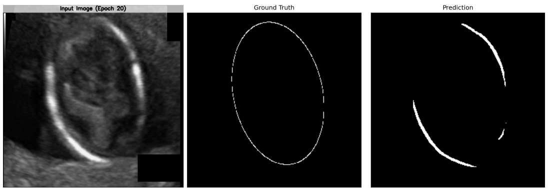

(a) UKAN - Epoch 10



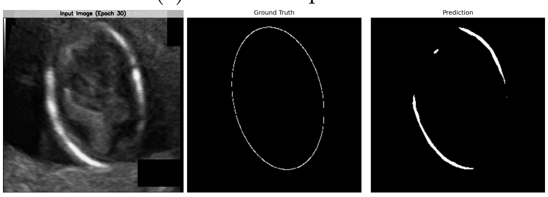

(b) Upgraded UKAN - Epoch 10



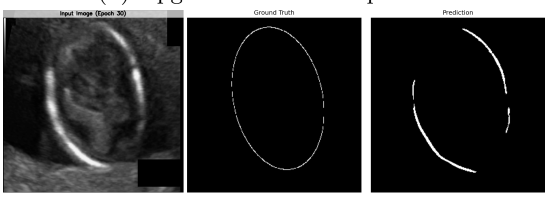

(c) UKAN - Epoch 20




(d) Upgraded UKAN - Epoch 20




(e) UKAN - Epoch 30




(f) Upgraded UKAN - Epoch 30

Figure 4: Comparison of predictions between the normal and upgraded models at different epochs.

### 4.2.3 Model Comparison

Table 2: Model Performance Comparison - Fetus Head Segmentation Dataset

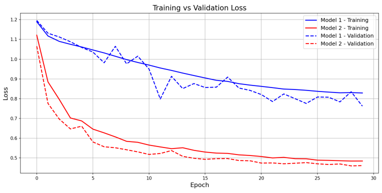
| Metric | Model 1 (Best) | Model 2 (Best) | Improvement | Relative Improvement (%) |
|---|---|---|---|---|
| Loss | 1.0203 | 0.8033 | 0.2170 | 21.26 |
| IoU | 0.1400 | 0.1647 | 0.0246 | 17.58 |
| F1 Score | 0.2469 | 0.2828 | 0.0359 | 14.54 |
| Val Loss | 1.0224 | 0.8232 | 0.1991 | 19.48 |
| Val IoU | 0.1348 | 0.1481 | 0.0133 | 9.83 |
| Val Dice | 0.2373 | 0.2573 | 0.0200 | 8.44 |

Model 1: UKAN Baseline, Model 2: Upgraded UKAN
Negative values indicate performance degradation

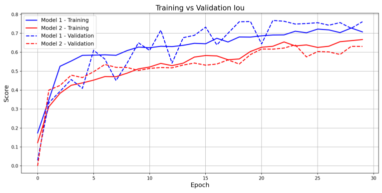
## 4.3 Lumbar Spine Segmentation Dataset

### 4.3.1 Training Progress

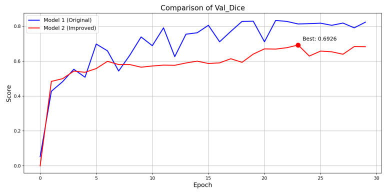
The training progress for the models is depicted through the following plots: loss, IoU, and DICE score over 30 epochs.




(a) Loss over epochs




(b) IoU over epochs

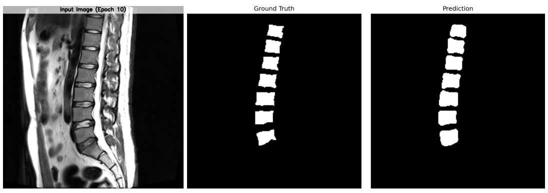



(c) DICE score over epochs

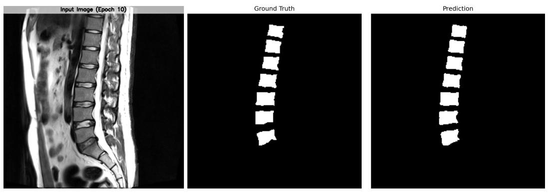
Figure 5: Training progress plots for the Lumbar Spine Segmentation dataset.

### 4.3.2 Epoch-wise Predictions

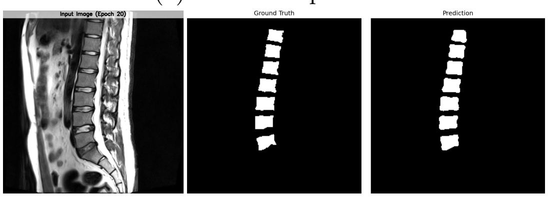
The following images show the model's predictions at different epochs, illustrating the improvement in segmentation accuracy over time. The following images show the model's predictions at different epochs, comparing the normal model (left) with the upgraded model (right).



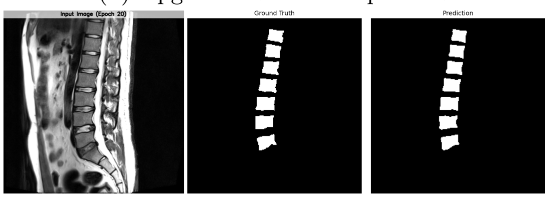

(a) UKAN - Epoch 10



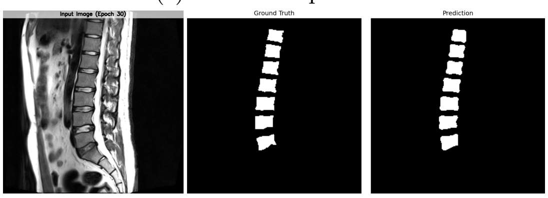

(b) Upgraded UKAN - Epoch 10



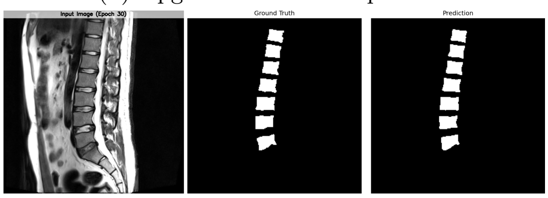

(c) UKAN - Epoch 20




(d) Upgraded UKAN - Epoch 20




(e) UKAN - Epoch 30




(f) Upgraded UKAN - Epoch 30

Figure 6: Comparison of predictions between the normal and upgraded models at different epochs.

### 4.3.3 Model Comparison

Table 3: Model Performance Comparison - Lumbar Spine Segmentation Dataset

| Metric | Model 1 (Best) | Model 2 (Best) | Improvement | Relative Improvement (%) |
|---|---|---|---|---|
| Loss | 0.8289 | 0.4842 | 0.3447 | 41.59 |
| IoU | 0.7278 | 0.6671 | −0.0608 | −8.35 |
| F1 Score | 0.8488 | 0.8135 | −0.0353 | −4.16 |
| Val Loss | 0.7623 | 0.4598 | 0.3026 | 39.69 |
| Val IoU | 0.7680 | 0.6392 | −0.1288 | −16.77 |
| Val Dice | 0.8336 | 0.6926 | −0.1410 | −16.92 |

Model 1: UKAN Baseline, Model 2: Upgraded UKAN
Negative values indicate performance degradation

# 5 Discussion

This study evaluated an **Upgraded UKAN** (Model 2), featuring CLAHE preprocessing and a Transformer bottleneck, against a baseline UKAN (Model 1) on ISIC 2017, Fetus Head, and Lumbar Spine datasets. The results demonstrated **highly dataset-dependent efficacy** of the enhancements:

- **ISIC 2017 (Table 1):** Mixed outcomes - Model 2 reduced training/validation losses by 55.4%/40.5% yet showed slight validation metric declines (-0.4% IoU, -0.8% Dice). This loss-accuracy divergence suggests overfitting or suboptimal task alignment despite better optimization.
- **Fetus Head (Table 2):** Consistent gains - 9.8%/8.4% relative improvements in validation IoU/Dice confirm architectural benefits. CLAHE-enhanced contrast and transformer context modeling likely enhanced fetal structure delineation in ultrasound images.
- **Lumbar Spine (Table 3):** Revealed a critical paradox - while Model 2 achieved a 41.59% loss reduction (0.8289 $\rightarrow$ 0.4842), validation metrics severely degraded (-16.77% IoU, -16.92% Dice). This exposes fundamental limitations in slice-based processing of 3D medical data:
  - CLAHE's slice-wise contrast enhancement disrupts cross-slice intensity coherence vital for spinal structures
  - Transformer attention fails to model inter-slice anatomical relationships critical for 3D segmentation
  - Loss minimization becomes decoupled from clinical accuracy metrics in volumetric contexts

# 6 Conclusion

The UKAN architecture demonstrates **modality-specific effectiveness** with critical lessons:

- **2D Dominance:** Achieved 55.4% loss reduction (ISIC) and 9.8% IoU gains (Fetus), validating CLAHE-transformers for slice-based analysis
- **3D Paradox:** Despite 3D medical inputs, slice processing degraded spine segmentation (-16.8% IoU, 16.9% Dice $\downarrow$), exposing volumetric modeling needs
- **Architectural Tradeoffs:**
  - Transformer-KAN synergy improves optimization but risks metric divergence (39.7% loss $\downarrow$ vs. -16.8% IoU)
  - CLAHE requires dataset-specific tuning (effective for dermis/fetus, harmful for spine)
- **Implementation Imperatives:**
  - True 3D processing with axial-attention transformers
  - Anatomically-constrained loss functions
  - Multi-scale CLAHE adaptation